PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A61K 48/00, C07K 14/435, C12N 1/21, 5/10, 15/12, 15/63, C12P 21/02

A1

(11) International Publication Number: WO 95/17213

(43) International Publication Date: 29 June 1995 (29.06.95)

(21) International Application Number: PCT/US94/14665

(22) International Filing Date: 21 December 1994 (21.12.94)

(30) Priority Data:
08/171,118 21 December 1993 (21.12.93) US

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(71) Applicant: SLOAN-KETTERING INSTITUTE FOR CANCER RESEARCH [US/US]; 1275 York Avenue, New York, NY 10021 (US).

(72) Inventor: PAVLETICH, Nikola, P.; Apartment 9C, 303 East 60th Street, New York, NY 10021 (US).

(74) Agent: WHITE, John, P.; Cooper & Dunham, 1185 Avenue of the Americas, New York, NY 10036 (US).

Express Mail No. EJ676053101US
PD-5780-01-TMC

(54) Title: P53-BASED POLYPEPTIDE FRAGMENTS, NUCLEIC ACID MOLECULES ENCODING SAME, AND USES THEREOI

(57) Abstract

The subject invention provides a nucleic acid molecule which encodes a polypeptide comprising a portion of p53 protein, which polypeptide is (a) capable of specifically binding to DNA having the sequence specifically recognized by p53 protein, (b) capable of specifically activating transcription of the DNA sequences whose transcription is activated by p53, (c) incapable of tetramerizing, and (d) incapable of binding to MDM-2 protein, and the polypeptide encoded thereby. The subject invention also provides a vector comprising the recombinant nucleic acid molecule of the subject invention. The subject invention further provides a host vector system for the production of a polypeptide comprising a portion of p53 protein. The subject invention further provides a method for producing the polypeptide of the subject invention using the host vector system of the subject invention. The subject invention further provides a pharmaceutical composition which comprises an effective amount of a recombinant virus capable of infecting a suitable host cell, said recombinant virus comprising a nucleic acid molecule encoding the polypeptide of the subject invention and capable of being expressed in the suitable host cell, and a pharmaceutically acceptable carrier.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

p53-BASED POLYPEPTIDE FRAGMENTS, NUCLEIC ACID MOLECULES ENCODING SAME, AND USES THEREOF

This invention was made with support under Grant No. GM 31471 from the National Institutes of Health. Accordingly, the U.S. government has certain rights in the invention.

Background of the Invention

Throughout this application, various publications are referenced by Arabic numerals. Full citations for these references may be found at the end of the specification immediately preceding the claims. The disclosure of these publications is hereby incorporated by reference into this application to describe more fully the art to which this invention pertains.

The p53 tumor suppressor gene is the most commonly mutated gene identified in human cancers, and mounting evidence points to the inactivation of p53 as a critical step leading to neoplastic transformation (reviewed in 1-3). Loss of p53 function results in an enhanced frequency of genomic rearrangements (4, 5), and eliminates the growth arrest response induced by DNA damage (6, 7). These studies suggest that p53 controls a cell cycle checkpoint that is important for maintaining the integrity of the genome (8, 9).

The exact mechanisms through which p53 exerts its biological function are not known, but its ability to bind to specific DNA sequences (10-13) and activate transcription (14-18), and its interactions with several cellular proteins, such as the MDM-2 oncogene product (19, 20), seem to play important

roles. For example, tumor-derived p53 mutants which have lost their ability to cause cell cycle arrest also are inactive in DNA binding and transactivation (16, 18). Furthermore, many of these mutants apparently promote neoplastic processes by forming hetero-oligomers with wild type p53 and abrogating its activity in a dominant negative manner (21, 22). The transcriptional activation function of p53 has been mapped to residues 1-42 at the N-terminus (14-18), and the oligomerization activity has been roughly mapped to the C-terminal portion (22, 23). It has been proposed that the DNA-binding domain may also reside in the C-terminal region of p53 (24), but no data has been presented to show that this region is involved in sequence-specific DNA binding.

In the experiments described hereinbelow, proteolytic digestion was used to identify the major structural domains in the human p53 protein, and these domains were then expressed in Escherichia coli and their functions characterized in vitro. The proteolytic digestion experiments show that the N-terminal portion of p53 is highly susceptible to proteolytic digestion and suggest that it is solvent-exposed and loosely folded. In contrast, the highly conserved middle portion of p53 is strikingly resistant to proteolytic digestion and thus seems to be an independently folded, compact structural domain. Finally, the pattern of cleavage sites in the C-terminal portion of p53 reveals a second, smaller structural domain. In vitro experiments using the recombinant domains show that the central, core domain contains the sequence-specific DNA binding activity of p53, and this activity appears to be dependent on zinc binding. The C-terminal domain contains the tetramerization activity of p53. Contrary to a previous report (23), a 20 residue basic region does not seem to be necessary for tetramer formation.

In a subject suffering from a neoplasm associated with the presence of a mutant p53 protein, the introduction of native p53 into the subject (e.g., through gene therapy) would not be expected to ameliorate the effects of the mutant p53 protein. Specifically, native p53, when introduced into the neoplastic cells of the subject, would (a) tetramerize with the mutant p53 protein via the C-terminal domain and (b) bind to the MDM-2 oncogene protein product via the N-terminal domain, thus preventing the native p53 protein from exerting its positive effect on the neoplastic cell through site-specific DNA binding and transcription activation.

The subject invention provides, inter alia, a polypeptide comprising a portion of p53 protein, which polypeptide is (a) capable of specifically binding to DNA having the sequence specifically recognized by p53 protein, (b) capable of specifically activating transcription of the DNA sequences whose transcription is activated by p53, (c) incapable of tetramerizing, and (d) incapable of binding to MDM-2 protein. The subject invention also provides related nucleic acid molecules, a pharmaceutical composition comprising same, and a method for using the pharmaceutical composition to treat a subject suffering from a neoplasm associated with the presence of mutant p53 protein in the cells of the subject. Accordingly, the subject invention overcomes the above-identified problem resulting from introducing native p53 into the neoplastic cells of a subject suffering from a neoplasm associated with the presence of mutant p53 protein.

Summary of the Invention

The subject invention provides a recombinant nucleic acid molecule which encodes a polypeptide comprising a portion of p53 protein, which polypeptide is (a) capable of specifically binding to DNA having the sequence specifically recognized by p53 protein, (b) capable of specifically activating transcription of the DNA sequences whose transcription is activated by p53, (c) incapable of tetramerizing, and (d) incapable of binding to MDM-2 protein.

The subject invention further provides the polypeptide encoded by the recombinant nucleic acid molecule of the subject invention.

The subject invention also provides a vector comprising the recombinant nucleic acid molecule of the subject invention.

The subject invention further provides a host vector system for the production of a polypeptide comprising a portion of p53 protein, which polypeptide is (a) capable of specifically binding to DNA having the sequence specifically recognized by p53 protein, (b) capable of specifically activating transcription of the DNA sequences whose transcription is activated by p53, (c) incapable of tetramerizing, and (d) incapable of binding to MDM-2 protein, which comprises the vector of the subject invention in a suitable host.

The subject invention further provides a method for producing a polypeptide comprising a portion of p53 protein, which polypeptide is (a) capable of specifically binding to DNA having the sequence specifically recognized by p53

protein, (b) capable of specifically activating transcription of the DNA sequences whose transcription is activated by p53, (c) incapable of tetramerizing, and (d) incapable of binding to MDM-2 protein, which comprises growing the host vector system of the subject invention under conditions permitting the production of the polypeptide and recovering the polypeptide produced thereby.

The subject invention further provides a pharmaceutical composition which comprises an effective amount of a recombinant virus capable of infecting a suitable host cell, said recombinant virus comprising a nucleic acid molecule encoding the polypeptide of the subject invention and capable of being expressed in the suitable host cell, and a pharmaceutically acceptable carrier.

Finally, the subject invention provides a method for treating a subject suffering from a neoplasm associated with the presence of mutant p53 protein in the cells of the subject, which comprises administering to the subject an amount of the pharmaceutical composition of the subject invention effective to treat the subject.

Brief Description of the Figures

Figure 1A

Digestion of human p53 by subtilisin. Amino acid sequence of the human p53 protein showing the subtilisin cutting points (residues C-terminal to the cleaved peptide bonds) identified in this study are shown. The sequences of the protease-resistant core domain and the C-terminal domain are underlined.

Figure 1B

Digestion of human p53 by subtilisin. SDS gel (12-20% polyacrylamide gradient stained with silver showing a titration of the reaction of subtilisin with p53. On the right half of the figure, the results of extensive microchemical and mass spectroscopic analyses are summarized; for lack of space, the data are not presented herein. Reactions contain the following ratios of subtilisin to p53 (weight:weight): lane 1, no subtilisin; lane 2, 10%; lane 3, 25%; lane 4, 60%; lane 5, 200%, lane 6, 500%. Lane 7 contains a reaction similar to that of lane 6, but has no p53. Some of the fragments give broader bands on the gel, and it is believed that this may reflect heterogeneity in the precise cleaving points. The C-terminal fragments are not readily apparent in this gel because of their low mass relative to the rest of the fragments.

Figure 2A

Comparison of the DNA binding activities of the core domain and of intact p53. Gel mobility shift experiments using the RGC site. Reactions in lanes 2-6 contain 1.1 μM of purified core domain peptide, and those in lanes 7-11 contain 0.2 μM of immunopurified human p53. Lane 1 contains no protein.

As sequence-specific competitors, either the RGC-binding site or a mutated RGC site was used; lanes 2 and 7, no competitor; lanes 3 and 8, 100 ng RGC site; lanes 4 and 9, 1000 ng mutant RGC site. Binding reactions contained 25 mM Tris-CL, pH 7.2, 175 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT, 5% Glycerol, 50 μg/ml bovine serum albumin, 28 μg/ml of nonspecific DNA, and 1-2 nM radiolabelled RGC site DNA.

Figure 2B

Comparison of the DNA binding activities of the core domain and of intact p53. Methylation of critical guanines in the RGC site interferes with core domain binding. The two lanes on the left show the piperidine cleavage pattern of the methylated upper strand. B; core domain-bound DNA recovered from the gel. F; free DNA recovered from the gel. The two lanes on the right show the pattern obtained with the lower strand. Dots represent the methylation-sensitive guanines of the core domain-RGC complex. For comparison, the methylation-sensitive guanines of the intact p53-RGC complex (10) are indicated on the RGC sequence: shadowed guanines show strong interference, and outlined guanines show weak interference. The locations of the pentamer sequence motifs are indicated by arrows. Solid arrows show the pentamers that match the consensus, and the dashed arrow shows the pentamer that has two mismatches.

Figure 3A

Competition experiments showing that the core domain requires two pentamer motifs for sequence-specific binding. Binding reactions, similar to those of Figure 2A, contained the core domain peptide (1.2 μM) and labelled RGC site (1 ng). The core domain-RGC complex was challenged with increasing amounts (10 ng, 33 ng, 100 ng, 333 ng and 1000 ng) of the unlabelled competitor fragments indicated on the

top of the figure. Lanes 1-5, a site with no pentamer motifs (but containing the same flanking sequences as the sites with pentamers); lanes 6-10, a site with a single pentamer; lanes 11-15, a site with two pentamers; lanes 16-20, a site with four pentamers; lanes 21-25, the RGC site; lanes 26-30, the mutant RGC site.

Figure 3B

Competition experiments showing that the core domain requires two pentamer motifs for sequence-specific binding. The results of the competition experiments were quantitated, and the amount of binding observed was plotted against the logarithm (base 10) of the amount of competitor used in each reaction. To facilitate comparisons, the amount of each competitor required to give a 50 percent reduction in binding is listed. In calculating the molar excess of the pentamer competitors, the fourth, imperfect pentamer of the RGC site is included.

Figure 4

Gel mobility shift assay showing that DNA binding by the core domain and by intact p53 requires metal binding. Binding reactions, similar to those of Figure 2A, with either core domain (lanes 1-4) or with intact p53 (lanes 5-8) contained the following amounts of 1,10-phenanthroline: lanes 1 and 5, no 1,10-phenanthroline; lanes 2 and 6, 0.1 mM; lanes 3 and 7 1 mM; lanes 4 and 8, 10 mM.

Figure 5A

Comparison of the oligomerization activities of the C-terminal domain and the C-terminal domain plus basic region peptides. Elution profiles of the two peptides on a Superdex 75 gel filtration column. The position of standards used to estimate the molecular weights are indicated on top. No other peaks are observed over a

concentration range of 1 μM - 1 mM for either peptide.

Figure 5A

Comparison of the oligomerization activities of the C-terminal domain and the C-terminal domain plus basic region peptides. Products of the reaction of glutaraldehyde with p53[311-365] and with p53[311-393] were separated on an SDS gel (15% polyacrylamide), and visualized with Coomasie blue staining. The oligomerization states of the various products were estimated from their mobilities relative to those of the molecular weight markers.

Figure 6

Structural domains of p53. Boxes with Roman numerals indicate the five regions of p53 that are conserved across species, and the bar graph above shows the approximate position and frequency of tumor-derived mutations (1). The positions of the DNA-binding domain (residues 102-292) and the tetramerization domain (residues 311-365) are indicated below.

Detailed Description of the Invention

The subject invention provides a recombinant nucleic acid molecule which encodes a polypeptide comprising a portion of p53 protein, which polypeptide is (a) capable of specifically binding to DNA having the sequence specifically recognized by p53 protein, (b) capable of specifically activating transcription of the DNA sequences whose transcription is activated by p53, (c) incapable of tetramerizing, and (d) incapable of binding to MDM-2 protein.

As used herein, a recombinant nucleic acid molecule is a nucleic acid molecule which does not occur in nature and which is obtained through the use of recombinant nucleic acid technology.

As used herein, p53 protein means full length, naturally occurring human p53 protein. The term "p53 protein" is herein used synonymously with the terms "p53", "native p53," "wild-type p53" and "intact p53." The amino acid and nucleic acid sequences of p53 are known in the art, and may be found, inter alia, at GenBank (Los Alamos, NM).

As used herein, "capable of specifically binding to DNA having the sequence specifically recognized by p53 protein" means capable of binding to DNA having the sequence specifically recognized by p53 protein, but incapable of binding to DNA having another sequence. The sequence specifically recognized by p53 protein includes, by way of example, the RGC site described infra.

As used herein, "capable of specifically activating transcription of the DNA sequences whose transcription is

activated by p53" means capable of increasing the rate of transcription of only certain regions of the cellular genome, said certain regions being the regions whose rate of transcription is increased by p53.

In one embodiment, the portion of the polypeptide of the subject invention responsible for specifically activating transcription of the DNA sequences whose transcription is activated by p53 comprises a portion of the p53 domain responsible therefor. The portion of p53 may further comprise either a point mutation or a deletion mutation which mutation renders that portion capable of specifically activating transcription but incapable of binding MDM-2 protein. In another embodiment, the portion of the polypeptide of the subject invention responsible for specifically activating transcription of the DNA sequences whose transcription is activated by p53 comprises a portion of a protein other than p53. By way of example, the protein other than p53 may be Herpes Simplex virus protein vp16.

In the preferred embodiment, the portion of the polypeptide of the subject invention responsible for specifically activating transcription of the DNA sequences whose transcription is activated by p53 is proximal to the amino terminus of the polypeptide relevant to the portion capable of specifically binding to DNA having the sequence specifically recognized by p53 protein.

As used herein, "incapable of tetramerizing" means incapable of forming a tetramer, i.e., a four-membered oligomer, with either p53 protein or mutant p53 protein in any permutation thereof.

In one embodiment, the recombinant nucleic acid molecule is a DNA molecule. The DNA molecule may be a cDNA molecule.

In another embodiment, the recombinant nucleic acid molecule is an RNA molecule. The RNA molecule may be an mRNA molecule.

In one embodiment, the portion of p53 protein has an N-terminus selected from the group consisting of about amino acid residues +94, +96 and +102, and a C-terminus selected from the group consisting of about amino acid residues +292, +305 and +310. In the preferred embodiment, the portion of p53 protein comprises the portion from about amino acid residue +102 to about amino acid residue +292.

The subject invention further provides the polypeptide encoded by the recombinant nucleic acid molecule of the subject invention. In one embodiment, the polypeptide encoded by the recombinant nucleic acid molecule of the subject invention is a purified polypeptide.

The subject invention also provides a vector comprising the recombinant nucleic acid molecule of the subject invention. Vectors include, by way of example, pVL1392 (for use in the construction of recombinant baculovirus) and pET3d expression vector. In one embodiment, the vector is a plasmid. In another embodiment, the vector is a virus.

In accordance with the invention, numerous vectors for the expression of the polypeptide of the subject invention may be employed. For example, one class of vectors utilizes DNA elements which are derived from animal viruses such as bovine papilloma virus, polyoma virus, adenovirus, vaccinia virus, retroviruses (RSV, MMTV or MoMLV), Semliki Forest virus or SV40 virus. Additionally, cells which have stably integrated the DNA into their chromosomes may be selected by introducing one or more markers which allow for the selection of transfected host cells. The marker may

provide, for example, prototropy to an auxotrophic host, biocide resistance, (e.g., antibiotics) or resistance to heavy metals such as copper or the like. The selectable marker gene can be either directly linked to the DNA sequences to be expressed, or introduced into the same cell by cotransformation. Additional elements may also be needed for optimal synthesis of mRNA. These elements may include splice signals, as well as transcriptional promoters, enhancers, and termination signals.

The subject invention further provides a host vector system for the production of a polypeptide comprising a portion of p53 protein, which polypeptide is (a) capable of specifically binding to DNA having the sequence specifically recognized by p53 protein, (b) capable of specifically activating transcription of the DNA sequences whose transcription is activated by p53, (c) incapable of tetramerizing, and (d) incapable of binding to MDM-2 protein, which comprises the vector of the subject invention in a suitable host. Methods of making host vector systems are well known to those skilled in the art.

In one embodiment, the suitable host is a bacterial cell. In the preferred embodiment, the bacterial cell is an E. coli cell. In another embodiment, the suitable host is an eucaryotic cell.

The subject invention further provides a method for producing a polypeptide comprising a portion of p53 protein, which polypeptide is (a) capable of specifically binding to DNA having the sequence specifically recognized by p53 protein, (b) capable of specifically activating transcription of the DNA sequences whose transcription is activated by p53, (c) incapable of tetramerizing, and (d) incapable of binding to MDM-2 protein, which comprises

growing the host vector system of the subject invention under conditions permitting the production of the polypeptide and recovering the polypeptide produced thereby.

Methods and conditions for growing host vector systems and for recovering the polypeptides so produced are well known to those skilled in the art, and may be varied or optimized depending upon the specific vector and host cell employed. Such recovery methods include, by way of example, gel electrophoresis, ion exchange chromatography, affinity chromatography or combinations thereof.

The subject invention further provides a pharmaceutical composition which comprises an effective amount of a recombinant virus capable of infecting a suitable host cell, said recombinant virus comprising a nucleic acid molecule encoding the polypeptide of the subject invention and capable of being expressed in the suitable host cell, and a pharmaceutically acceptable carrier.

The effective amount of the recombinant virus may be determined according to methods well known to those skilled in the art. Methods of making and using recombinant viruses are well known to those skilled in the art. A suitable host cell includes a neoplastic cell whose neoplastic state is associated with the presence therein of mutant p53 protein.

Pharmaceutically acceptable carriers are well known to those skilled in the art and include, but are not limited to, 0.01-0.1M and preferably 0.05M phosphate buffer or 0.8% saline. Additionally, such pharmaceutically acceptable carriers may be aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic

esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers such as those based on Ringer's dextrose, and the like. Preservatives and other additives may also be present, such as, for example, antimicrobials, antioxidants, chelating agents, inert gases and the like.

In one embodiment, the recombinant virus is a retrovirus and the nucleic acid molecule is an RNA molecule. Retroviruses include any RNA virus that uses reverse transcriptase during replication and is capable of incorporating its genome into the host cell genome (e.g., Rous Sarcoma virus, Mouse Mammary Tumor virus and HIV). Methods of making and using recombinant retroviruses are well known to those skilled in the art.

The subject invention further provides a method for treating a subject suffering from a neoplasm associated with the presence of mutant p53 protein in the cells of the subject, which comprises administering to the subject an amount of the pharmaceutical composition of the subject invention effective to treat the subject. In the preferred embodiment, the subject is a human.

The amount of the pharmaceutical composition of the subject invention effective to treat the subject may be determined according to methods well known to those skilled in the art.

Neoplasms associated with the presence of mutant p53 protein include, by way of example, osteosarcomas and neoplasms of

the lung, brain, liver, esophagus, bladder and ovary. A more exhaustive list of such neoplasms is provide by Hollstein, et al. (1).

Finally, the subject invention provides a method for determining the three dimensional structure of the DNA-binding domain of p53 which comprises the steps of (a) producing crystals of a portion of p53 protein capable of specifically binding to DNA having the sequence specifically recognized by p53 protein, and (b) analyzing the crystals so produced using X-ray crystallographic methods so as to thereby determine the three-dimensional structure of the DNA-binding domain of p53.

This invention will be better understood by reference to the Experimental Details which follow, but those skilled in the art will readily appreciate that the specific experiments detailed are only illustrative of the invention as described more fully in the claims which follow thereafter.

Experimental Details

I - Introduction

Mutations in the p53 tumor suppressor gene are the most commonly observed genetic alterations in human cancer. The majority of these mutations occur in the conserved central portion of the gene, but there has been little information about the function of this region. Using proteolytic digestion of the 393 amino acid human p53 protein, a 191 amino acid protease-resistant fragment (residues 102-292) was identified which corresponds to the central portion of p53, and it is this core fragment which is the sequence-specific DNA-binding domain of the protein. DNA-binding is inhibited by metal chelating agents, and it was found that the core domain contains zinc. Proteolytic digests also reveal a fifty-three amino acid C-terminal domain which is shown to be the tetramerization domain of p53.

II - Materials and Methods

A - Proteolytic digestion and identification of the fragments

Proteolysis reactions contained 60-200 μg/ml of human p53 protein immunopurified from Sf21 cells expressing the recombinant baculovirus pEV55hwt (25), in a buffer of 10 mM Hepes-Na, pH 7.5, 5 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 50% Glycerol, and the protease subtilisin (the p53 preparations used in this study typically contain a number of minor contaminant proteins that have not been characterized). The reactions were typically carried out for 25 minutes and stopped by 1 mM PMSF. For analysis, the fragments were either purified by reversed phase HPLC on a C4 column, or separated by SDS PAGE and transferred onto Immobilon-P

membrane (Millipore). Amino acid analyses were performed on an ABI 420H amino acid analyzer with PITC derivation, and N-terminal sequencing was performed on an ABI 477A pulsed liquid microsequencer with on line PTH analysis. The mass of the purified fragments was determined either on an electrospray ionization triple-stage quadropole mass spectrometer (Finnigan), or on a matrix-assisted laser desorbtion time of flight mass spectrometer (Vestek or Finnigan). It should be noted that although most of the major fragments have been identified, there are several amino proteolytic fragments that have not been characterized.

B - Expression and purification of the core domain

The portion of the human p53 gene encoding residues 102-292 was amplified in a polymerase chain reaction, and the amplified product was cloned into the pET3d expression vector (Novagen). The core domain peptide was produced in *E. coli* BL21(D3) cells at room temperature. cells were harvested by centrifugation and lysed by sonication in 40 mM Mes-Na, pH 6.0, 200 mM NaCl, and 5 mM DTT. The lysate was clarified by centrifugation, diluted 5-fold, loaded onto a Mono S cation exchange column (Pharmacia) in 40 mM Mes-Na, pH 6.0, 5 mM DTT, and was eluted by a NaCl gradient to yield the core domain at >85% purity. For most of the experiments, the Mono S fraction was precipitated by 80% ammonium sulfate and was further purified by gel filtration chromatography on a Superdex 75 gel filtration column (Pharmacia), in 50 mM Bistrispropane-Na, pH 6.8, 100 mM NaCl, 5 mM DTT, to yield the core domain at >98% purity.

C - Gel mobility shift and methylation interference assays

Binding reactions (15 μl) contained 25 mM Tris-Cl, pH 7.2, 175 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT, 5% Glycerol, 50 μg/ml

bovine serum albumin, 28 μg/ml of nonspecific DNA, and 1-2 nM radiolabelled RGC site DNA. The reactions were equilibrated at room temperature for 20 minutes, and were electrophoresed at 10 V/cm in a 4.5% polyacrylamide gel running 0.5x Tris-borate buffer. Typically, 50% binding was obtained with reactions that contained 8.5 μg/ml of intact p53 or 23 μg/ml of core domain. For the competition experiments, the radiolabelled RGC site was first mixed with the appropriate unlabelled DNA and then was equilibrated with the protein. The amount of bound and free DNA was quantitated using phosphorimager (Fuji). The methylation interference analysis was performed essentially as described (Current Protocols in Molecular Biology, Wiley Interscience, NY). Briefly, the core domain-RGC complex was formed under conditions described earlier, except the protein concentration was adjusted to give approximately 30% bound DNA. The free and core domain-bound RGC fragments were separated on a 4.5% polyacrylamide gel and the DNA was eluted in 0.5 M ammonium acetate and 1 mM EDTA at 50°C. The DNA was cleaved at the methylated sites using piperidine, and equal amounts of labelled DNA were electrophoresed on a denaturing polyacrylamide (12%) sequencing gel.

D - <u>DNA fragments</u>

The DNA fragments used in the study were constructed by annealing purified synthetic oligonucleotides. The sequences of these fragments, reading along one strand, are as follows (the pentamer sequences are indicated in bold letters): RGC (42bp): GATCCGATTGCCTTGCCTGGACTTGCCTGGC CTTGCCTTTTG; mutant RGC (42 bp): GATCCGATTCCCTTCCGTGCAGTTCC GTGGCCTTCCCTTTTG; [The wild-type RGC site contains nine guanines which appear to be important for binding as demonstrated by methylation interference experiments (10), and the mutant RGC site that has six of these guanines mutated to cytosines.] nonspecific DNA (38 bp): CGTACTTATC

GAGCGGGGGCGTAGTGATAGTTCTCTAG; no-pentamer (36 bp): GATCCGTATTCTCTTTTCCTAATTACAATTCGATTG; one pentamer (41 bp): GATCCGTATTCTCTTTTCCTAGACAAATTACAATTCGATTG; two pentamers (46 bp): GATCCGTATTCTCTTTTCCTAGACATGCCTAATTACAATTCGATTG; four pentamers (56 bp): GATCCGTATTCTCTTTTCCTTGACTTGCCCAGACATGTTTAATTACAATTCGATTG.

E - Plasma emission spectroscopy

Plasma emission spectroscopy was performed using a 31 element simultaneous, inductively coupled plasma (EPA method 6010). A typical analysis used a 1.5 ml solution of 46 μM core domain peptide (concentration was determined by amino acid composition analysis) in gel filtration buffer, and yielded 1.4 parts per million (21.4 μM) zinc. This analysis has a detection limit for zinc of 0.05 ppm. The stoichiometry corresponds to 0.47 moles of zinc per mole of core domain. An analysis of buffer showed that there were no contaminating metals.

F - Expression, purification and crosslinking of the C-terminal peptides

The cloning and expression of the C-terminal peptides was done according to the procedures used for the core domain peptide. The p53[311-365] peptide was purified as follows: the E. coli lysate was acidified with 100 mM Na-acetate, pH 4.6, clarified by centrifugation, loaded onto a Mono S cation exchange column running 100 mM Na-acetate, pH 4.6, and was eluted with a NaCl gradient. The p53[311-393] peptide was extracted from the insoluble fraction of the E. coli lysate using 6.4 M Guanidine-Cl, and was purified by reversed phase HPLC on a C4 column (the mobile phase contained 0.1% trifluoroacetic acid, and the peptide was eluted with an acetonitrile gradient). Gel filtration experiments with the C-terminal peptides were performed using a Superdex 75 (Pharmacia) column running a buffer of

100 mM Tris-Cl, pH 7.6, and 200 mM NaCl. No changes in the elution profiles were observed over a concentration range of 1 μM - 1 mM of either peptide. Crosslinking reactions contained 167 μM of peptide and EM grade glutaraldehyde (Sigma) in 50 mM Hepes-Na, pH 8.0, and 150 mM NaCl. After 30 minutes, the reactions were stopped by the addition of 100 mM Tris-Cl.

III - **Results**

A - Proteolytic digestion of human p53 reveals two structural domains and a loosely folded N-terminal region

The proteolytic digestion experiments were done with immunopurified human p53 (25) and the protease subtilisin. Subtilisin has relatively low sequence specificity, making it a useful probe of structural domains. The subtilisin concentration was varied over a wide range to determine which regions of the molecule are most and least susceptible to proteolysis. After subtilisin digestion, the proteolytic fragments were purified, and were analyzed by N-terminal sequencing, mass spectroscopy, and amino acid analysis. The cutting points inferred from these analyses are presented in Figure 1A.

Figure 1B shows the pattern of fragments produced when human p53 is digested with increasing concentrations of subtilisin. Cutting occurs most readily at the extreme N-terminal (at residues 8, 18, 23) and at three clusters of sites in the C-terminal portion of p53 (at residues 305 and 311; at residues 364, 366 and 368; and at residues 387 and 388). The digestion pattern in the C-terminal region reveals the existence of a fifty-three amino acid fragment (residues 311-364) which is relatively resistant to proteolytic digestion, and suggests that this fragment is an

independently folded structural domain (Figure 1B, lanes 2 and 3). This C-terminal domain is very readily cleaved from p53 and it is presumed that it is connected to the rest of the protein by a flexible linker. Immediately after this domain, and separated from it by a protease-sensitive linker, there is a roughly 20 amino acid region (residues 368-387) rich in basic amino acids. It appears that this basic region has considerable secondary structure since subtilisin does not readily cut internal to it, even though it cuts the flanking regions with ease.

Digestion with higher subtilisin concentrations gives additional cuts in the N-terminal region, which is cleaved progressively at residues 39, 40, 44, 54 and 68, and eventually at residues 94, 96 and 102. The observed digestion pattern indicates that the region containing residues 1-67 is readily accessible to subtilisin, suggesting that it is generally solvent-exposed and loosely folded. The transcriptional activation function of p53 has been attributed to residues 1-42 in this region (16), and the accessibility or flexibility of this region may be important for the protein-protein interactions involved in transcriptional activation.

At high concentrations, subtilisin also cuts the region between residues 293 and 306, giving a 191 amino acid core fragment (residues 102-292) that is strikingly resistant to digestion and persists even at a five-fold excess (weight:weight) of subtilisin (Figure 1B, lane 6). This clearly shows that the core fragment is an independently folded, compact structural domain. Consistent with this evidence for a tightly folded structure, the core domain contains the evolutionarily highly conserved regions of p53. These include residues 117-142, 171-181, 234-258, and 270-286. The one conserved region of p53 that is not in the

core domains is found near the N-terminus region, at residues 13-19. The core domain also contains the majority of the mutation sites identified in tumors which are concentrated in the conserved regions in the central portion of p53 (1).

B - The core domain contains the sequence-specific DNA-binding activity of p53

Using the gel mobility shift assay, the products of the subtilisin digestion reaction were tested for DNA binding and it was found that the fragments retained significant activity (data not shown). Binding to a p53 site could be observed even when the digestion mixture contained no detectable intact p53 and consisted predominantly of the core domain. The shifted DNA band had a mobility faster than that of the intact p53-DNA complex, suggesting that a proteolytic fragment of p53, presumably the core domain, was responsible for the DNA binding activity. The subtilisin digest contained small amounts of other fragments as well. Thus, to test this more carefully, the core domain (residues 102-292) was expressed in E. coli, purified to near homogeneity, and then subjected to gel mobility-shift experiments with a p53-binding site from the ribosomal gene cluster promoter (RGC site; 10).

Figure 2A shows that the recombinant core domain can form a complex with the RGC site and gives a shifted band with a mobility faster than that of the intact p53-RGC site complex (in the presence of 175 mM NaCl and 28 μM intact p53 or with 1.1 μM core domain). To determine whether the core domain binds specifically to the RGC site, the complex was challenged with either the wild type RGC site or a mutant site. Figure 2A shows that, like intact p53 (lanes 8-11), the core domain strongly prefers the wild type RGC site over the mutant site (lanes 3-6).

The DNA sequence specificities of the core domain and of intact p53 were also compared by means of methylation interference assay. Kern et al. (10) reported that the RGC site contains several guanines which, when methylated by DMS, interfere with the binding of p53. The methylation interference pattern observed with the core domain-RGC complex (Figure 2B) reveals that this complex is also sensitive to guanine methylation. When this pattern is compared to the pattern observed with intact p53 (as reported by Kern et al. (10)) it appears that the methylation-sensitive guanines of the core domain-RGC complex form a subset of those found with the intact p53-complex. The methylated guanines that interfere most strongly with the binding of intact p53 also interfere with the binding of the core domain (these are guanine 25 on the upper strand and guanines 20', 21' and 25' on the lower strand; Figure 2B). On the other hand, the methylated guanines that only partially interfere with the binding of intact p53 (10) show little or no interference with the binding of the core domain (these are guanines 15, 19, and 20 on the upper strand and guanines 30' and 31' on the lower strand; Figure 2B). These results demonstrate that the core domain can make many of the key DNA contacts made by the intact p53 protein.

The methylation interference patterns of the core domain and intact p53 can also be interpreted by referring to the pentamer consensus Pu-Pu-Pu-C-(A/T). The majority of the known p53-binding sites contain four repeats of this pentamer motif (although some of the repeats only partially match the consensus). The RGC site contains three repeats with a perfect match to the consensus, and one repeat with a partial match (Figure 2B). In the case of intact p53, the methylated guanines that interfere most strongly with binding occur in only two of these repeats; the third repeat

shows only weak interference, and the fourth (imperfect) repeat shows no interference at all. As mentioned above, the core domain contacts mimic the strongest contacts seen with the intact protein. The methylated guanines that interfere strongly also occur in the first two repeats, whereas the third and fourth repeats do not show any significant interference. In summary, methylation interference shows that the core domain interacts preferentially with two of the pentamer motifs of the RGC site, and these are the same regions where the intact p53 makes its strongest contacts.

C - <u>Two pentamer sequences are required for the binding of the core domain</u>

p53 has been shown to form tetramers (26), and since binding sites typically contain four pentamer sequence motifs it seems possible that each p53 molecule interacts with a single pentamer motif. This model would predict that the isolated core domain, which is missing the oligomerization domain and appears to be a monomer in solution (gel filtration data; not shown), might bind to a single pentamer sequence. However, methylation interference experiments suggest that the situation is more complicated, and additional experiments were done to determine how many pentamers are required for tight binding by the core domain. Competition experiments were used similar to those of Figure 2A, and the core domain-RGC complex was challenged with sites that contained one, two, or four pentamer motifs. Figure 3A shows that the most effective competitors are the RGC site and a four pentamer site. A site with two pentamers is also a good competitor, while a site with a single pentamer is a poor competitor (comparable to the mutant RGC site or to a site with no pentamer sequences). When the amount of competitors required to give 50% reduction in binding are compared (Figure 3B), it becomes

apparent that the core domain requires at least two pentamer sequences for specific binding. This finding suggests that two core domain molecules may interact with two pentamers in a cooperative fashion. Interestingly, when the amount of competitor is normalized for the pentamer content (Figure 3B), the two pentamer site seems to have the same affinity for the core domain as the four pentamer site. Although cooperative binding (to form a dimer on the DNA) seems to be the most plausible explanation for the methylation interference and competition data, one cannot rule out the alternative possibility that one core domain molecule may interact simultaneously with two adjacent pentamers.

D - <u>Zinc is required for the DNA binding activity</u>

The p53 protein contains 10 cysteines--all of them in the core domain--and this raises the possibility that p53 binds to a metal ion. To address this possibility, the effect of metal chelating agents on DNA binding was tested. Figure 4 shows that the metal chelating agent 1,10-phenanthroline abolishes DNA binding by both the core domain and intact p53. Similar results were obtained with EDTA (data not shown). The metal seems to be tightly bound to p53 since relatively high concentrations of metal chelating reagents are needed to abolish DNA binding (approximately 5 - 10mM of 1,10-phenanthroline or EDTA).

To analyze the metal content of the core domain, plasma emission spectroscopy was performed, which analysis can detect 31 different elements including Ca, Co, CU, Fe, Mg, Mn, and Zn. It was found that the only metal which the core domain contains is zinc. This result, taken together with the finding that metal chelating agents abolish DNA binding, suggest that p53 is a zinc metalloprotein.

The plasma emission analysis was repeated several times, and

the zinc content consistently corresponded to a ratio of about 0.5 moles of zinc per mole of core domain. Although this stoichiometry could in principle result from a metal bridged dimer, it is unlikely since the core domain elutes from a gel filtration column as a monomer. It seems more likely that some of the zinc dissociates from the protein during purification, since it was found that the core domain is sensitive to oxidation.

E - <u>The C-terminal structural domain contains the tetramerization activity of p53</u>

Previous studies (23, 22) had mapped the oligomerization activity of p53 to the C-terminal portion, and Stürzbecher et al. have reported that a stretch of basic residues in this region is required for the conversion of dimers to tetramers. The proteolytic digestion pattern observed indicates the presence of a fifty-three residue structural domain (residues 311-365) in this C-terminal region of the protein. The short basic region implicated in oligomerization (residues 368-387) occurs immediately after this domain and is connected to it by a linker highly sensitive to subtilisin.

To determine whether the fifty-three residue structural domain has any oligomerization activity, and to investigate the role of the basic region in oligomerization, the structural domain (residues 311-365) and a peptide consisting of this domain plus the basic region (residues 311-393) were expressed in <u>E. coli</u>, and the oligomerization activity of the purified peptides was assayed. Each peptide elutes from a Superdex 75 gel filtration column as a single, high molecular weight complex, and the two peptides seem to have similar oligomerization states (Figure 5A): the shorter peptide (residues 311-365) elutes at an estimated molecular weight of 37 kDa (about 6.4 molecules/ oligomer) and the

longer peptide (residues 311-393) at 53 kDa (about 5.9 molecule/oligomer). Since the elution profile on a gel filtration column may be influenced by the molecular shape, glutaraldehyde crosslinking experiments were also performed to determine whether these oligomeric complexes might actually be tetramers (Figure 5B). Crosslinking of either peptide produces a ladder of four bands with sizes roughly corresponding to monomers, dimer, trimers, and tetramers. Crosslinking of the longer peptide is more efficient, presumably because it contains 11 lysine residues whereas the shorter peptide only contains 5 lysines. Since the ladder ends with the tetramer band, and no pentamers, hexamers, or other higher order oligomers are observed, this suggests that the tetramer is the predominant form of these peptides. These results, taken together with the results of the gel filtration experiments, demonstrate that the C-terminal structural domain is the tetramerization domain of p53, and show that the basic region is not required for tetramerization. It should be noted that crosslinking of the longer peptide also produces high molecular weight (>85kDa) products (Figure 5B, lane 4). Although these may represent nonspecific crosslinking due to the high lysine content of the basic region, it is also conceivable that the basic region is involved in the association of tetramers to form higher order oligomers observed with intact p53 (26). However, such high molecular weight forms were not observed in the gel filtration experiments (Figure 5A), and the significance of these bands is not clear.

Since it has been proposed that the DNA-binding domain of p53 may reside in the C-terminus (24), the C-terminal peptides were also tested for DNA binding. It was found that the longer peptide (residues 311-393) has considerable affinity for DNA, but the shorter peptide (residues 311-365) which is missing the basic region has no detectable affinity

for DNA (data not shown). However, the interaction of the longer peptide with DNA is not specific for the p53-binding site, since it was found that the mutant RGC site, or nonspecific calf thymus DNA can effectively compete with the wild type RGC site for binding. Although the affinity of the C-terminal peptide for DNA may be an in vitro artifact, it is also possible that the basic region makes auxiliary contact as p53 binds to DNA, or binds nonspecifically and helps regulate p53's activity (27).

IV - Discussion

The proteolytic digestion experiments and the studies of fragments produced in E. coli provide new insights into the structural and functional organization of the p53 protein. The key finding of this study is that the central portion of p53 (residues 102-292) constitutes the sequence-specific DNA-binding domain of the protein. This finding reveals that the DNA-binding domain coincides with the major mutation hotspots (Figure 6), and this helps us understand why the tumor-derived mutant p53 proteins are defective in DNA binding.

Several lines of evidence demonstrate that the core domain is the sequence-specific DNA-binding domain of p53. First, the core domain binds tightly to the RGC site, with an affinity comparable to that of intact p53. Second, competition experiments show that the core domain, like intact p53, strongly prefers to bind to the wild type RGC site over the mutant RGC site. Third, methylation interference experiments show that the core domain makes many of the key DNA contacts made by intact p53. The methylated guanines that interfere most strongly with the binding of intact p53 are precisely the ones that interfere with the binding of the core domain. Fourth, the pentamer

sequence motif that is required for the binding of intact p53 is also necessary for the binding of the core domain.

Although the core domain appears to be a monomer in solution, several observations suggest that it may bind as a dimer. A monomer might be expected to bind weakly, but it was found that the core domain binds to the RGC site with an affinity only five to six-fold lower than that of intact p53. It was also found that the core domain requires at least two adjacent pentamers for specific binding. Again these results suggest that the core domain may form dimers upon DNA binding and thus recognize the two pentamers in a cooperative fashion.

In addition to the effects of oligomerization, there are likely to be other effects that modulate the DNA binding activity of the intact protein. Hupp, et al. (27) have shown that p53's DNA binding activity may be regulated by phosphorylation in the C-terminus, and it is conceivable that there may be additional regulatory mechanisms.

The proteolytic digestion experiments reveal that p53 also contains a smaller structural domain located in the C-terminal region (residues 311-363). Following this fifty-three residue domain, there is a flexible linker, and then a twenty residue highly basic region (residues 368-387). These studies of peptides from this region show that the shorter C-terminal domain is the tetramerization domain of p53, and the basic region seems to be involved in nonspecific DNA binding. These finding are in general agreement with previous data that mapped the location of the oligomerization activity to the C-terminal portion of p53 (22, 23). However, some aspects of the oligomerization model of Sturzbecker et al. (23) are inconsistent with these findings. They proposed that the region corresponding to

the shorter C-terminal domain is involved in dimerization and that the basic region is required to convert the dimers to tetramers. Several lines of evidence suggest that the basic region is not required for tetramer formation. First, a peptide containing the C-terminal domain alone elutes from a gel filtration column in a high molecular weight form, having essentially the same oligomerization state as the C-terminal domain plus basic region peptide. Second, glutaraldehyde crosslinking experiments demonstrate that the preferred oligomerization state of both peptides is the tetramer. Third, Shaunlian et al. (22) clearly show that a peptide corresponding to the C-terminal domain (residues 302-360) has transforming properties very similar to those of a peptide containing the basic region (residues 302-390), suggesting that the two peptides have very similar oligomerization activities. The transforming ability of these peptides seems to involve the formation of mixed oligomers with wild type p53.

These results, in conjunction with previous work from other groups, allow a clear delineation of the major domains of p53. The N-terminal region contains the transactivation domain, i.e., the transcription activating domain; the central region contains the sequence-specific DNA-binding domain; and the C-terminal region contains sites responsible for oligomerization and nonspecific DNA binding. The results presented herein enable an understanding of the role of the conserved central region and the effects of mutations in this region. The identification of p53's structural domains also should aid in its structural characterization by crystallographic or NMR spectroscopic methods.

References

1. Hollstein, M., D. Sidransky, B. Vogelstein, and C.C. Harris. 1991. p53 mutations in human cancers. Science 253:49-53.

2. Levine, A.J., J. Momand, C.A. Finlay. 1991. The p53 tumor suppressor gene. Nature 351:453-456.

3. Vogelstein, B. K.W. Kinzler. 1992. p53 function and dysfunction. Cell 70:523-526.

4. Livingstone, L.R., A. White, J. Sprouse, E. Livanos, T. Jacks and T.D. Tlsty. 1992. Altered cell cycle and gene amplification potential accompany loss of wild-type p53. Cell 70:923-935.

5. Yin, Y., M.A. Tainsky, F.Z. Bischoff, L.C. Strong, and G.M. Wahl. 1992. Wild-type p53 restores cell cycle control and inhibits gene amplification in cells with mutant p53 alleles. Cell 70:937-948.

6. Kuerbitz, S.J., B.S. Plunkett, W.V. Walsh, M.B. Kastan. 1992. Wild-type p53 is a cell cycle checkpoint determinant following irradiation. Proc. Natl. Acad. Sci. U.S.A. 89:7491-7495.

7. Kastan, M.B., Q. Zhan, W.S. El-Deiry, F. Carrier, T. Jacks, W.V. Walsh, B.S. Plunkett, B. Vogelstein, and A.J. Fornace, Jr. 1992. A mammalian cell cycle checkpoint utilizing p53 and GADD45 is defective in ataxia-telangiectasia. Cell 71:587-597.

8. Lane, D.P. 1992. p53, guardian of the genome. Nature

358:15-16.

9. Hartwell, L. 1992. Defects in a cell cycle checkpoint may be responsible for the genomic instability of cancer cells. Cell 71:543-546.

10. Kern, S.E., K.W. Kinzler, A. Bruskin, D. Jarosz, P. Friedman, C. Prives, and B. Vogelstein. 1991. Identification of p53 as a sequence specific DNA-binding protein. Science 252:1708-1711.

11. Bargonetti, J., P.N. Friedman, S.E. Kern, B. Vogelstein and C. Prives. 1991. Wild-type but not mutant p53 immunopurified proteins bind to sequences adjacent to the SV40 origin of replication. Cell 65:1083-1091.

12. El-Deiry, W.S., S.E. Kinzler, J.A. Pietenpol, K.W. Kinzler and B. Vogelstein. 1992. Definition of a consensus binding site for p53. Nature Genetics 1:45-49.

13. Funk, W.D., D.T. Pak, R.H. Karas, W.E. Wright and J.W. Shay. 1992. A transcriptionally active DNA-binding site for human p53 complexes. Mol. Cell. Biol. 12:2866-2871.

14. Raycroft, L., H. Wu, G. Lozano. 1990. Transcriptional activation by wild-type but not transforming mutants of the p53 anti-oncogene. Science 249:1049-1051.

15. Fields, S., S.K. Jang. 1990. Presence of a potent transcriptional activating sequence in the p53 protein. Science 249:1046-1048.

16. Unger, T., M.M. Nau, S. Segal, J.D. Minna. 1992. p53:

a transdominant regulator of transcription whose function is ablated by mutations occurring in human cancer. EMBO J. 11:1383-1390.

17. Farmer, G., J. Bargonetti, H. Zhu, P. Friedman, R. Prywes and C. Prives. 1992. Wild-type p53 activates transcription in vivo. Nature 358:83-86.

18. Kern, S.E., J.A. Pietenpol, S. Thiagalingam., A. Seymour, K.W. Kinzler and B. Vogelstein. 1992. Oncogenic forms of p53 inhibit p53-regulated gene expression. Science 256:827-830.

19. Momand, J., G.P. Zambetti, D.C. Olson, D. George and A.J. Levine. 1992. The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. Cell 69:1237.

20. Oliner, J.D., K.W. Kinzler, P.S. Meltzer, D.L. George and B. Vogelstein. 1992. Amplification of a gene encoding a p53 associated protein in human sarcomas. Nature 358:80-83.

21. Milner, J., E.A. Medcalf. 1991. Cotranslation of activated mutant p53 with wild type drives the wild-type p53 protein into the mutant conformation. Cell 65:765-774.

22. Shaulian, E., A. Zauberman, D. Ginsberg, M. Oren. 1992. Identification of a minimal transforming domain of p53: negative dominance through abrogation of sequence-specific DNA binding. Mol. Cell. Biol. 12:5581-5592.

23. Stûrzbecher, H.W., R. Brain, C. Addison, K. Rudge, M.

Remm, M. Grimaldi, E. Keenan and J.R. Jenkins. 1992. A C-terminal α-helix plus basic region motif is the major structural determinant of p53 tetramerization. Oncogene 7:1513-1523.

24. Foord, O.S., P. Bhattacharya, Z. Reich, V. Rotter. 1991. A DNA binding domain is contained in the C-terminus of wild type p53 protein. Nucleic Acids Res. 19:5191-5198.

25. Bargonetti, J., J. Reynisdottir, P.N. Friedman, C. Prives. 1992. Site-specific binding of wild-type p53 to cellular DNA is inhibited by SV40 T antigen and mutant p53. Genes Dev. 6:1886-1898.

26. Stenger, J., G.A. Mayr, K. Mann, P. Tegtmeyer. 1992. Formation of stable p53 homotetramers and multiples of tetramers. Mol. Carcinog. 5:102-106.

27. Hupp, T.R., D.W. Meek, C.A. Midgley, D.P. Lane. 1992. Regulation of the specific DNA binding function of p53. Cell 71:875-886.

What is claimed is:

1. A recombinant nucleic acid molecule which encodes a polypeptide comprising a portion of p53 protein, which polypeptide is (a) capable of specifically binding to DNA having the sequence specifically recognized by p53 protein, (b) capable of specifically activating transcription of the DNA sequences whose transcription is activated by p53, (c) incapable of tetramerizing, and (d) incapable of binding to MDM-2 protein.

2. The recombinant nucleic acid molecule of claim 1, wherein the nucleic acid molecule is a DNA molecule.

3. The recombinant nucleic acid molecule of claim 2, wherein the DNA molecule is a cDNA molecule.

4. The recombinant nucleic acid molecule of claim 1, wherein the nucleic acid molecule is an RNA molecule.

5. The recombinant nucleic acid molecule of claim 1, wherein the portion of p53 protein has an N-terminus selected from the group consisting of about amino acid residues +94, +96 and +102, and a C-terminus selected from the group consisting of about amino acid residues +292, +305 and +310.

6. The recombinant nucleic acid molecule of claim 5, wherein the portion of p53 protein comprises the portion from about amino acid residue +102 to about amino acid residue +292.

7. The polypeptide encoded by the recombinant nucleic acid

molecule of claim 1.

8. A vector comprising the recombinant nucleic acid molecule of claim 3.

9. The vector of claim 8, wherein the vector is a plasmid.

10. The vector of claim 8, wherein the vector is a virus.

11. A host vector system for the production of a polypeptide comprising a portion of p53 protein, which polypeptide is (a) capable of specifically binding to DNA having the sequence specifically recognized by p53 protein, (b) capable of specifically activating transcription of the DNA sequences whose transcription is activated by p53, (c) incapable of tetramerizing, and (d) incapable of binding to MDM-2 protein, which comprises the vector of claim 8 in a suitable host.

12. The host vector system of claim 11, wherein the suitable host is a bacterial cell.

13. The host vector system of claim 12, wherein the bacterial cell is an _E. coli_ cell.

14. The host vector system of claim 11, wherein the suitable host is an eucaryotic cell.

15. A method for producing a polypeptide comprising a portion of p53 protein, which polypeptide is (a) capable of specifically binding to DNA having the sequence specifically recognized by p53 protein, (b) capable of specifically activating transcription of the DNA sequences whose transcription is activated by p53, (c) incapable of tetramerizing, and (d) incapable of

binding to MDM-2 protein, which comprises growing the host vector system of claim 11 under conditions permitting the production of the polypeptide and recovering the polypeptide produced thereby.

16. A pharmaceutical composition which comprises an effective amount of a recombinant virus capable of infecting a suitable host cell, said recombinant virus comprising a nucleic acid molecule encoding the polypeptide of claim 7 and capable of being expressed in the suitable host cell, and a pharmaceutically acceptable carrier.

17. The pharmaceutical composition of claim 16, wherein the recombinant virus is a retrovirus and the nucleic acid molecule is an RNA molecule.

18. A method for treating a subject suffering from a neoplasm associated with the presence of mutant p53 protein in the cells of the subject, which comprises administering to the subject an amount of the pharmaceutical composition of claim 16 effective to treat the subject.

19. The method of claim 18, wherein the subject is a human.

FIGURE 1A

```
              8                  18        23                            39 40     44
  1  MEEPQSDPSV EPPLSQETFS DLWKLLPENN VLSPLPSQAM DDLMLSPDDI
         54                 68                                   94  96
 51  EQWFTEDPGP DEAPRMPEAA PPVAPAPAAP TPAAPAPAPS WPLSSSVPSQ
     102
101  KTYQGSYGFR LGFLHSGTAK SVTCTYSPAL NKMFCQLAKT CPVQLWVDST
151  PPPGTRVRAM AIYKQSQHMT EVVRRCPHHE RCSDSDGLAP PQHLIRVEGN
201  LRVEYLDDRN TFRHSVVVPY EPPEVGSDCT TIHYNYMCNS SCMGGMNRRP
                                                 293
251  ILTIITLEDS SGNLLGRNSF EVRVCACPGR DRRTEEENLR KKGEPHHELP
         306      311
301  PGSTKRALPN NTSSSPQPKK KPLDGEYFTL QIRGRERFEM FRELNEALEL
                364 366 368                 387 388
351  KDAQAGKEPG GSRAHSSHLK SKKGQSTSRH KKLMFKTEGP DSD
```

FIGURE 1B

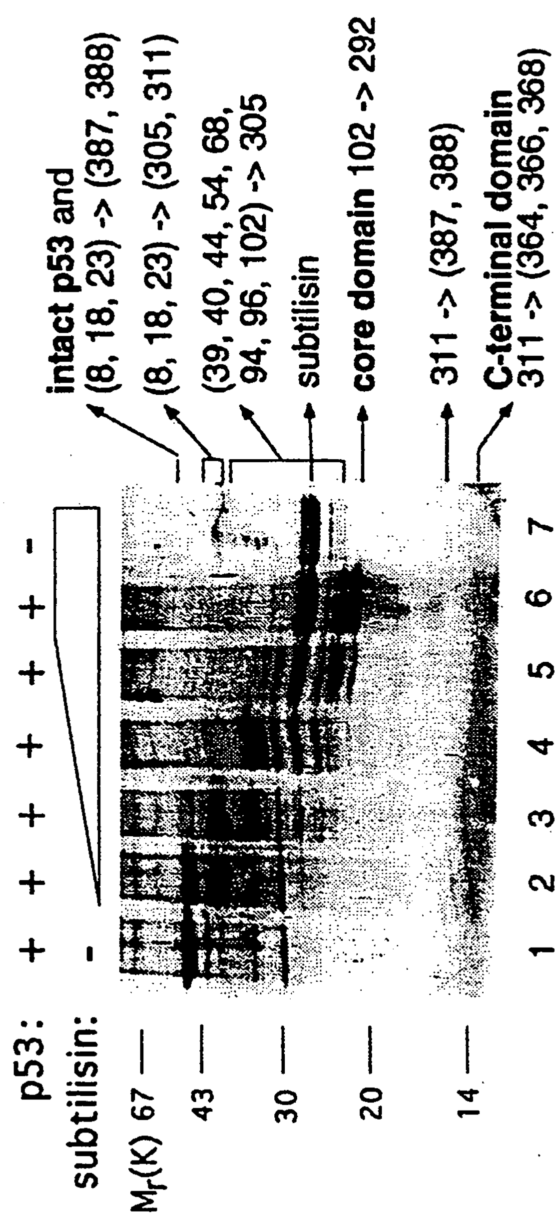

p53: + + + + + + -
subtilisin: -
M_r(K) 67
43
30
20
14
1 2 3 4 5 6 7
intact p53 and
(8, 18, 23) -> (387, 388)
(8, 18, 23) -> (305, 311)
(39, 40, 44, 54, 68,
94, 96, 102) -> 305
subtilisin
core domain 102 -> 292
311 -> (387, 388)
C-terminal domain
311 -> (364, 366, 368)

FIGURE 2A

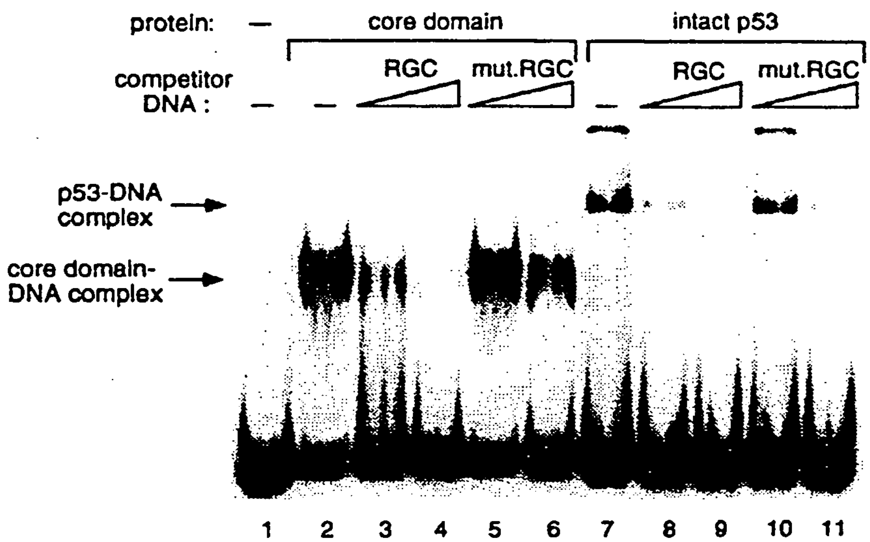

protein:
—
core domain
intact p53
competitor DNA :
—
—
RGC
mut.RGC
—
RGC
mut.RGC
p53-DNA complex
core domain-DNA complex
1 2 3 4 5 6 7 8 9 10 11

FIGURE 2B

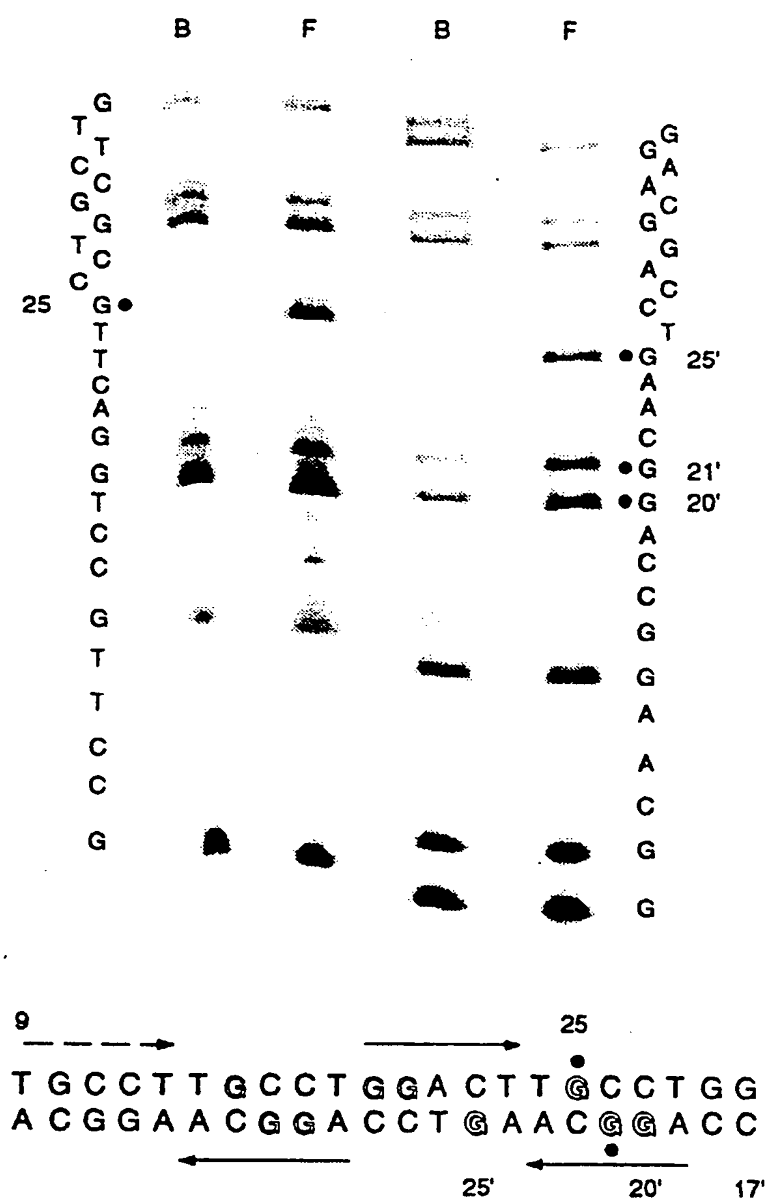

B F B F
25
25'
21'
20'
9
25
T G C C T T G C C T G G A C T T G C C T G G
A C G G A A C G G A C C T G A A C G G A C C
25'
20'
17'

FIGURE 3A

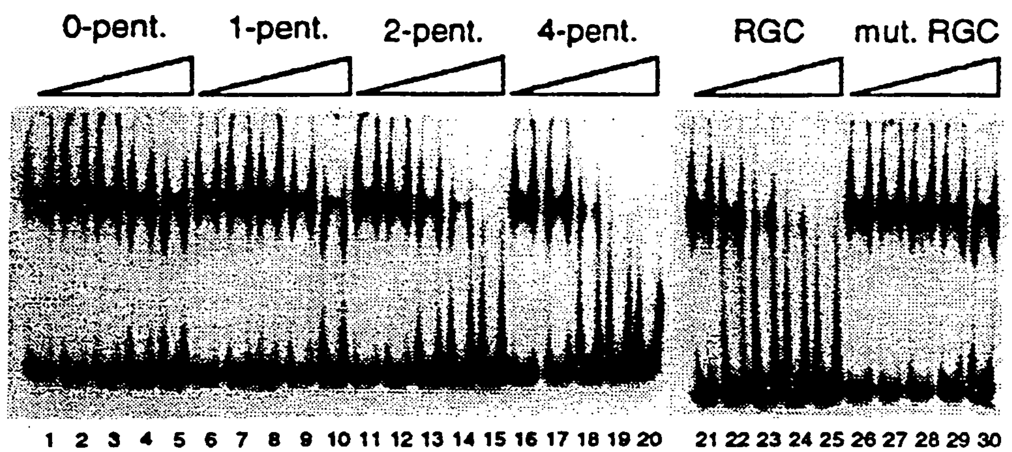
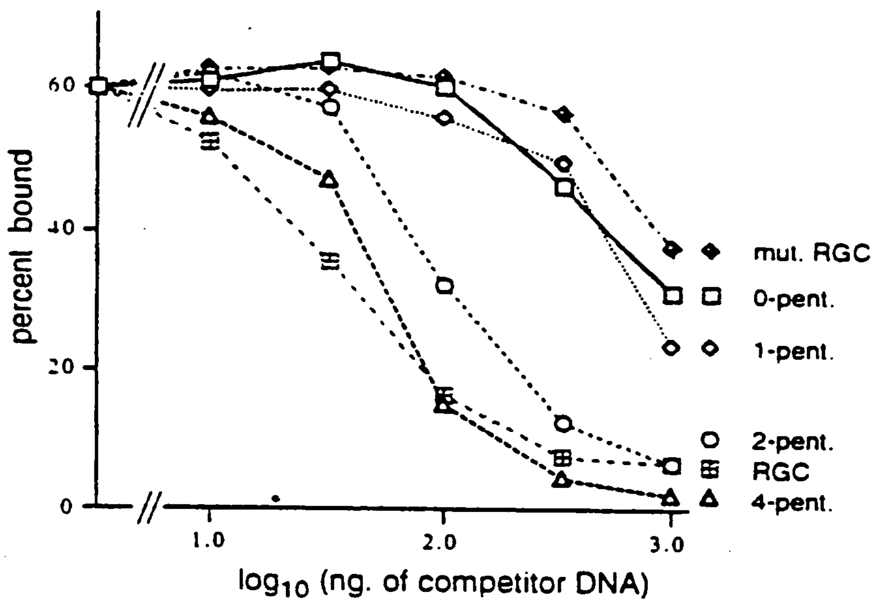

0-pent.
1-pent.
2-pent.
4-pent.
RGC
mut. RGC
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
21 22 23 24 25 26 27 28 29 30

FIGURE 3B



Excess of competitor that reduces binding by 50%

	By mass	By moles of pentamers *
mutant RGC †	1250	-
0-pentamer	1100	-
1-pentamer	800	200
2-pentamers	110	48
4-pentamers	60	43
RGC	50	50

† extrapolated
* corrected for length of fragment

FIGURE 4

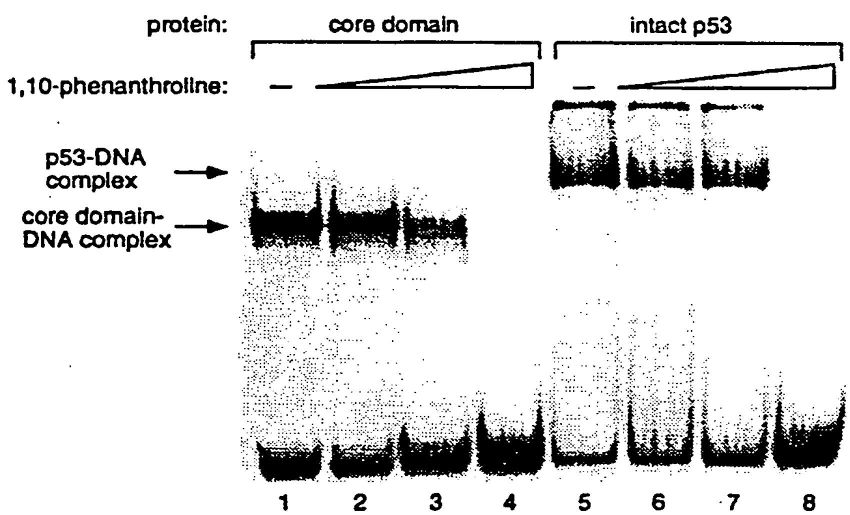

protein:
core domain
intact p53
1,10-phenanthroline:
p53-DNA complex
core domain-DNA complex
1
2
3
4
5
6
7
8

FIGURE 5A

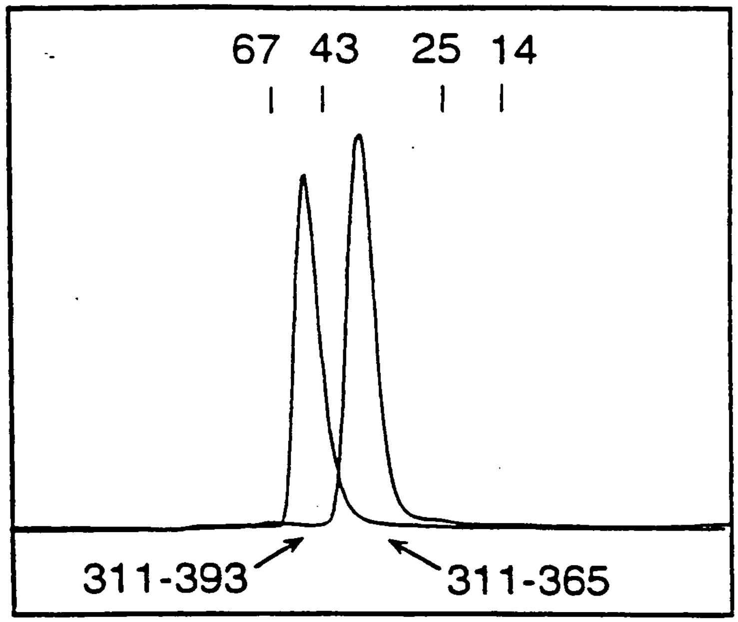


FIGURE 5B

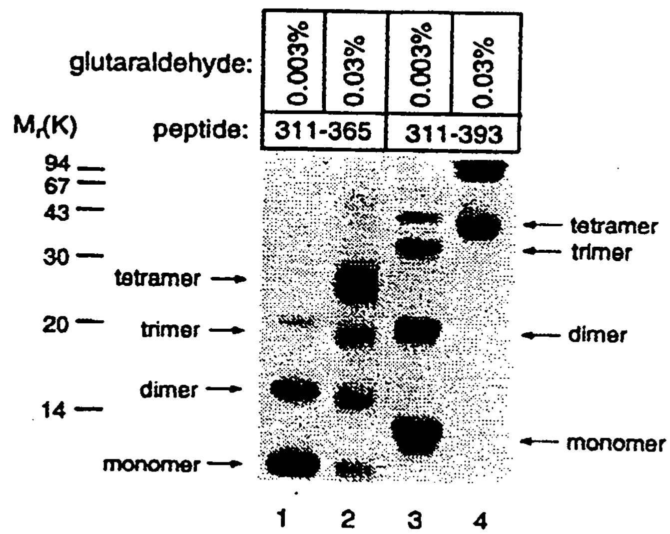
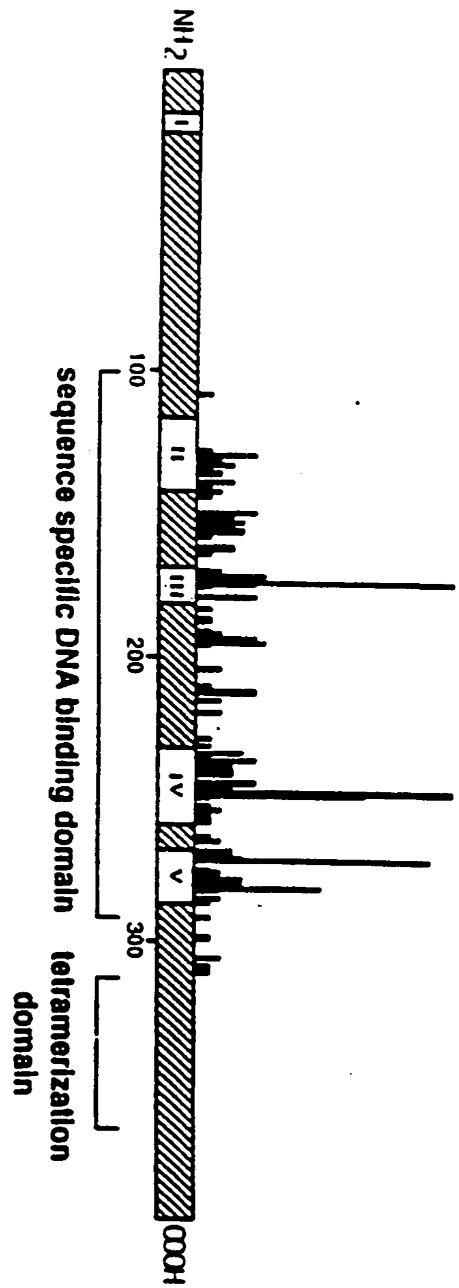

glutaraldehyde:
0.003%
0.03%
0.003%
0.03%
M_r(K)
peptide:
311-365
311-393
94
67
43
30
20
14
tetramer
trimer
dimer
monomer
tetramer
trimer
dimer
monomer
1
2
3
4

FIGURE 6


NH2
I
II
III
IV
V
100
200
300
sequence specific DNA binding domain
tetramerization domain
COOH

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/14665

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.1, 240.2, 252.3, 252.33, 320.1; 514/44; 530/350, 358; 536/23.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, Dialog
search terms: p53, tetramer?, delet?, domain, fragment, transcription

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Cell, Vol. 71, issued 27 November 1992, Hupp et al., "Regulation of the specific DNA binding function of p53", pages 875-886, see page 883.	15
A, P	Cell Growth & Differentiation, Vol. 5, issued July 1994, Zhang et al., "The DNA-binding and transcription-activation abilities of p53 are necessary but not sufficient for its antiproliferation function", pages 705-710, see the entire document.	1-19
A, P	Oncogene, Vol. 9, issued 1994, Marston et al., "Interaction of p53 with MDM2 is independent of E6 and does not mediate wild type transformation suppressor function", pages 2707-2716, see the entire document.	1-19

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search 13 MARCH 1995	Date of mailing of the international search report 27 MAR 1995
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer ERIC GRIMES Telephone No. (703) 308-0196

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Oncogene, Vol. 8, No. 11, issued 1993, Tarunina et al., "Human p53 binds DNA as a protein homodimer but monomeric variants retain full transcription transactivation activity", pages 3165-3173, see the entire document.	1-19
A, P	Science, Vol. 265, issued 15 July 1994, Cho et al., "Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations", pages 346-355, see the entire document.	1-19
A	Molecular and Cellular Biology, Vol. 13, No. 9, issued September 1993, Unger et al., "Functional domains of wild-type and mutant p53 proteins involved in transcriptional regulation, transdominant inhibition, and transformation suppression", pages 5186-5194, see the entire document.	1-19

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/14665

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

A61K 48/00; C07K 14/435; C12N 1/21, 5/10, 15/12, 15/63; C12P 21/02

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

435/69.1, 240.2, 252.3, 252.33, 320.1; 514/44; 530/350, 358; 536/23.5